

Prospects & Overviews

The phage-host arms race: Shaping the evolution of microbes

Adi Stern and Rotem Sorek*

Bacteria, the most abundant organisms on the planet, are outnumbered by a factor of 10 to 1 by phages that infect them. Faced with the rapid evolution and turnover of phage particles, bacteria have evolved various mechanisms to evade phage infection and killing, leading to an evolutionary arms race. The extensive co-evolution of both phage and host has resulted in considerable diversity on the part of both bacterial and phage defensive and offensive strategies. Here, we discuss the unique and common features of phage resistance mechanisms and their role in global biodiversity. The commonalities between defense mechanisms suggest avenues for the discovery of novel forms of these mechanisms based on their evolutionary traits.

Keywords:

■ arms race; bacteria; evolution; phage; resistance

Introduction

Phage-host relationships have been studied intensively since the early days of molecular biology. In the late 1970s, while viruses were found to be ubiquitous, it was assumed that they were present in relatively low numbers and that their effect on microbial communities was low [1]. With the increasing availability of new molecular techniques that allow studies of microbial communities without the need to culture them [2], it is now realized that viruses greatly outnumber bacteria in the ocean and other environments, with viral numbers ($\sim 10^7$ – 10^8 /mL) often 10-fold larger than bacterial cell counts ($\sim 10^6$ /mL) [3–5]. Thus, bacteria are confronted with a constant threat of phage predation.

The Red Queen hypothesis (Box 1) posits that competitive environmental interactions, such as those displayed by hosts and parasites, will lead to continuous variation and selection towards adaptation of the host, and counter-adaptations on the side of the parasite. Arguably, nowhere is this evolutionary trend so pronounced as in phage-microbe interactions. This is due to the extremely rapid evolution and turnover of phage particles [6], causing acute pressure on microbial communities to evade infection and killing by phages. In fact, the arms race between phages and bacteria is predicted to have had an impact on global nutrient cycling [7], on global climate [6, 7], on the evolution of the biosphere [8], and also on the evolution of virulence in human pathogens [9].

This review focuses on the evolution of three of the most well-studied microbial defense mechanisms against phages: the restriction-modification (RM) system, the recently discovered “clustered regularly interspersed palindromic repeats” (CRISPR) loci together with CRISPR-associated (*cas*) genes, and the abortive infection (Abi) system (summarized in Table 1). We first describe these defense systems, and the counter-adaptations that evolved in the phage to allow escape from bacterial defense. Next, we discuss features that are common to many microbial defense systems, such as rapid evolution, tendency for lateral gene transfer (LGT), and the selfish nature of these systems. We elaborate on defense systems that have gained new functions in the host genome. Finally, the exciting hypothesis that many other prokaryotic defense systems are still to be discovered is discussed. Our review mainly focuses on the evolutionary angle of the phage-

DOI 10.1002/bies.201000071

Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel

*Corresponding author:

Rotem Sorek

E-mail: rotem.sorek@weizmann.ac.il

Abbreviations:

Abi, abortive infection; **Cas**, CRISPR-associated sequence; **CRISPR**, clustered regularly interspersed palindromic repeats; **crRNA**, CRISPR RNA; **LGT**, lateral gene transfer; **MTase**, methyltransferase; **REase**, restriction endonuclease; **RM**, restriction modification; **TA**, toxin-antitoxin.

Box 1

The Red Queen hypothesis

The Red Queen hypothesis was originally proposed by Van Valen [10], and is also termed the evolutionary arms race hypothesis. As the Red Queen tells Alice in Lewis Carroll's *Through the Looking-Glass*: "Now, here, you see, it takes all the running you can do, to keep in the same place. If you want to get somewhere else, you must run at least twice as fast as that!" The original theory proposed that in tight co-evolutionary interactions, such as those in a prey-predator relationship, changes (e.g., running faster) on the one side may lead to near extinction of the other side. The only way the second side can maintain its fitness is by counter-adaptation (running even faster). This will lead to an uneasy balance between prey and predator, where species have to constantly evolve to stay at the same fitness level. The metaphor of an evolutionary arms race has been found to be relevant for many biological processes, but nowhere is this metaphor as apt as in host-parasite relationships.

host arms race; for deeper mechanistic descriptions of phage resistance systems we refer the readers to an excellent recent review [11]. In addition, the discussion here is restricted to active defense mechanisms; passive host adaptations, such as mutations at the phage receptors, are not discussed.

Restriction-modification systems: Degradation of foreign DNA

Arguably, the most well-studied phage defense mechanism is the RM system [12], which is present in over 90% of sequenced bacterial and archaeal genomes [13]. This system is composed of two activities: one that restricts incoming foreign genetic

material and another that protects host genetic material from restriction (Fig. 1A). Both activities are mediated by recognition of a specific DNA sequence, on average 4–8 bp long. Protection is normally conferred by modification (usually methylation) of specific bases in this recognition sequence in the host genome. Accordingly, all non-methylated DNA recognition sequences are recognized as foreign and are cleaved. Genetically, the minimal composition of RM systems consists of a methyltransferase (MTase) gene that performs the defense activity and a restriction endonuclease (REase) gene that performs the foreign restriction activity. Since the REases undergo rapid evolution, it is often the presence of an MTase that serves as the basis for identification of RM systems in newly sequenced genomes [14].

Phages have evolved to evade the ubiquitous RM systems in a variety of ways (Fig. 1B). Some phages have acquired an MTase, or stimulate the host MTase so that it confers protection to the phage genome [15]. Other phages code for proteins that target and shut down the REase. One interesting example is the Ocr protein of phage T7, which blocks the active site of some REases by mimicking 24 bp of bent B-form DNA [16]. Alternatively, some phages incorporate unusual bases in their genomes, thus throwing off the REase. For example, some *Bacillus subtilis* phages replace thymine with 5-hydroxymethyluracil [15]. These phages code a protein that then further inhibits the host protein uracil-DNA glycosylase from cleaving uracil bases from the phage DNA [17, 18]. Interestingly, this inhibitory protein is also a DNA mimic (reviewed in ref. 19). The T-even phages T2, T4, and T6 also contain unusual bases in their genomes and may further post-synthetically glycosylate their DNA to avoid REase restriction [15]. Yet another evasion mechanism that has been reported is alteration of the restriction recognition sites. For example, some phages employ "palindrome avoidance": since Type II REases often recognize symmetrical (palindromic) sequences, some phages tend to avoid containing such sites in their genome [20].

The well-studied model organisms *Escherichia coli* and T4 phage display a fascinating example of a co-evolutionary arms race [11]. The battle purportedly begins with the T4 phage genome, which contains the modified base hydroxymethylcytosine instead

Table 1. A summary of the major defense mechanisms described in this article

Name	Mechanism	Phylogenetic breadth
Restriction-modification	The restriction enzyme cleaves specific patterns in the incoming foreign DNA, while the modification enzyme protects host DNA from cleavage by unique biochemical modification	Appear in ~90% of all sequenced prokaryotic genomes [13]
CRISPR/Cas	Fragments of phage DNA are integrated into CRISPR loci, which are then transcribed and processed into short non-coding RNAs. These RNAs, along with the associated Cas proteins, guide the way to interfere with the phage nucleic acids, in an as-yet-unknown mechanism	Appear in ~40% and ~90% of all sequenced bacterial and archaeal genomes, respectively [27]. The distribution of CRISPR/Cas-bearing species across the phylogenetic tree of life is highly patchy [38, 59]
Abortive infection	Premature cellular death occurs upon phage entry, blocking the expansion of the phage to neighboring cells. Notably, Abi systems include a large collection of mechanisms with little or no known evolutionary relationship, apart from a very similar phenotype	Currently known Abi systems display a sporadic phylogenetic distribution in Gamma-proteobacteria and in Firmicutes [44], yet recent reports suggest that some systems may have an even broader phylogenetic range [54, 111]

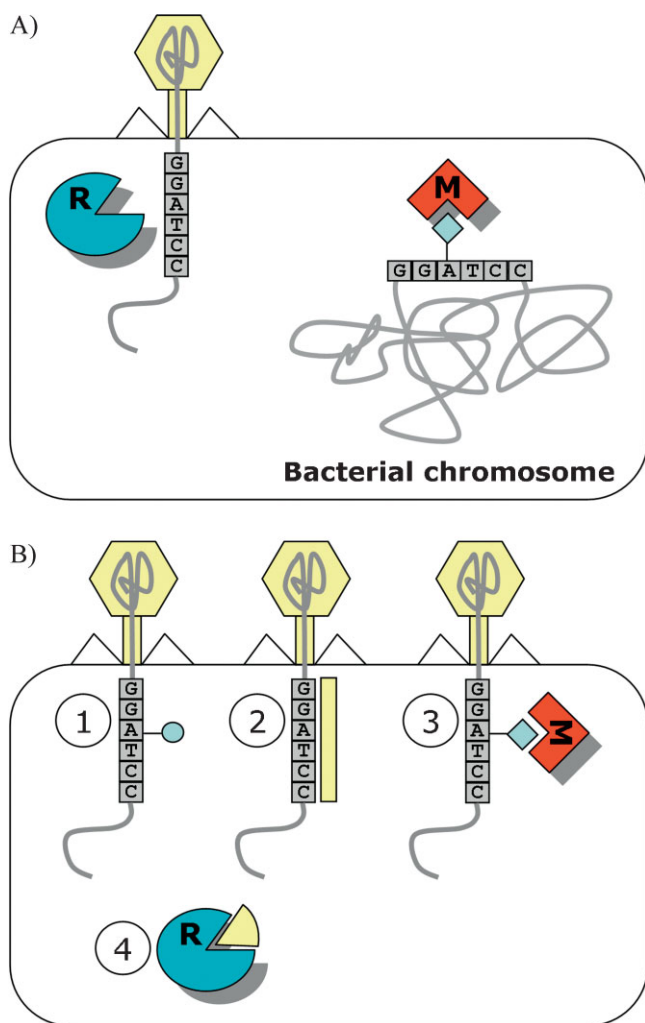


Figure 1. Restriction-modification (RM) defense system. **A:** A general illustration of function, exemplified by type II RM enzymes. **B:** Examples of strategies employed by phages to evade restriction: (1) incorporation of unusual bases protects from restriction [15]; (2) masking of the restriction sites by phage proteins [112]; (3) stimulation of MTase activity causes the phage DNA to be protected; (4) neutralization of REase by phage proteins that mimic DNA [113].

of cytosine [15]. To counter-attack the phage, *E. coli* K-12 possesses a unique form of REase, the McrBC enzyme, which cleaves only modified DNA substrates (such as that of T4) [21]. The T4 phage rises to the challenge by glucosylating its genome, and is thus impervious to McrBC [15]. However, *E. coli* cT596 encodes the GmrS-GmrD system that can also restrict glucosylated DNA [22, 23]. In continuation of the battle, some T4 phages encode the IPI protein, which in its processed form (IPI*) disables the GmrS-GmrD system [24]. Evidently, the battle is far from an end, since bacterial strains have been found to overcome the IPI* protein of T4 [24]. In fact, it is likely that many bacterial defense systems, coupled with their cognate phage evasion strategies, have undergone similar attack and counter-attack cycles, where a change on one side selects for changes that can overcome the opponent.

CRISPR/Cas: Acquired adaptive immunity

The existence of a unique nucleotide arrangement in *E. coli*, comprised of a cluster of direct repeats interspersed with variable sequences (termed spacers; both around 30 bp long), was first described in 1987 [25]. These clusters, together with associated *cas* genes (Fig. 2A), were later found to exist in ~40% of sequenced bacterial genomes and ~90% of archaeal genomes [26–28]. The first inkling that this system comprises a phage-defense mechanism arose when spacer sequences were found to be highly similar to DNA from foreign origin, *i.e.* from phage, plasmid or transposon DNA [29–31]. In 2007, experimental evidence was presented that this system indeed confers resistance to phage infection [32]. This led to the striking insight that similar to vertebrates, prokaryotes possess acquired (and inherited) immunity [27, 33–35]. Following phage infection, it has been found that a small portion of bacterial cells integrated new spacers identical to the phage genomic sequence (termed proto-spacer), resulting in CRISPR-mediated phage resistance [32, 36, 37]. Further experiments have shown that the CRISPR locus is transcribed into a single RNA transcript, which is then further cleaved by the Cas proteins to generate smaller CRISPR RNA (crRNA) units, each including one targeting spacer [37, 38]. These units then interfere with the incoming foreign genetic material by complementary base-pairing with either DNA [36, 37] or RNA [39] from the foreign element (Fig. 2A).

While the study of CRISPR/Cas is still in its infancy, examples of phages that are resistant to CRISPR/Cas interference have nevertheless been noted (Fig. 2B). Following the first round of infection and acquisition of novel spacers by the bacterial CRISPR, phages that have mutated, recombined or lost their proto-spacer target sequence in the second round of infection are then resistant to CRISPR [32, 40–42]. It also seems likely that phages have evolved mechanisms that directly target the CRISPR/Cas machinery. To date, only vague hints of such mechanisms exist. For example, it has been shown that one of the proteins encoded by the T7 phage phosphorylates the CasB protein [43]. It remains to be shown whether this feat of the phage affects CRISPR/Cas functioning.

Abortive infection: Cellular suicide

If a phage has successfully entered the host cell and avoided restriction by the host RM systems and by CRISPR, it proceeds to develop, replicate, and release its progeny. Abi is a collective term describing host mechanisms that interrupt with phage development at different stages of phage transcription, genome replication, and phage packaging [44]. Abi-mediated resistance leads to death (suicide) of the cell, and is thought to occur since corruption of host functions has already been initiated by the phage. However, this death confers an advantage to surrounding bacterial cells since it confines the infection to the sacrificed cell and prevents the spread of infectious particles.

Although several Abi systems have been discovered, with the majority of them encoded on plasmids, the mechanism by which they operate remains largely unknown. Frequently, Abi is mediated by a single gene encoded on a plasmid or on a prophage, which displays little or no homology to any known

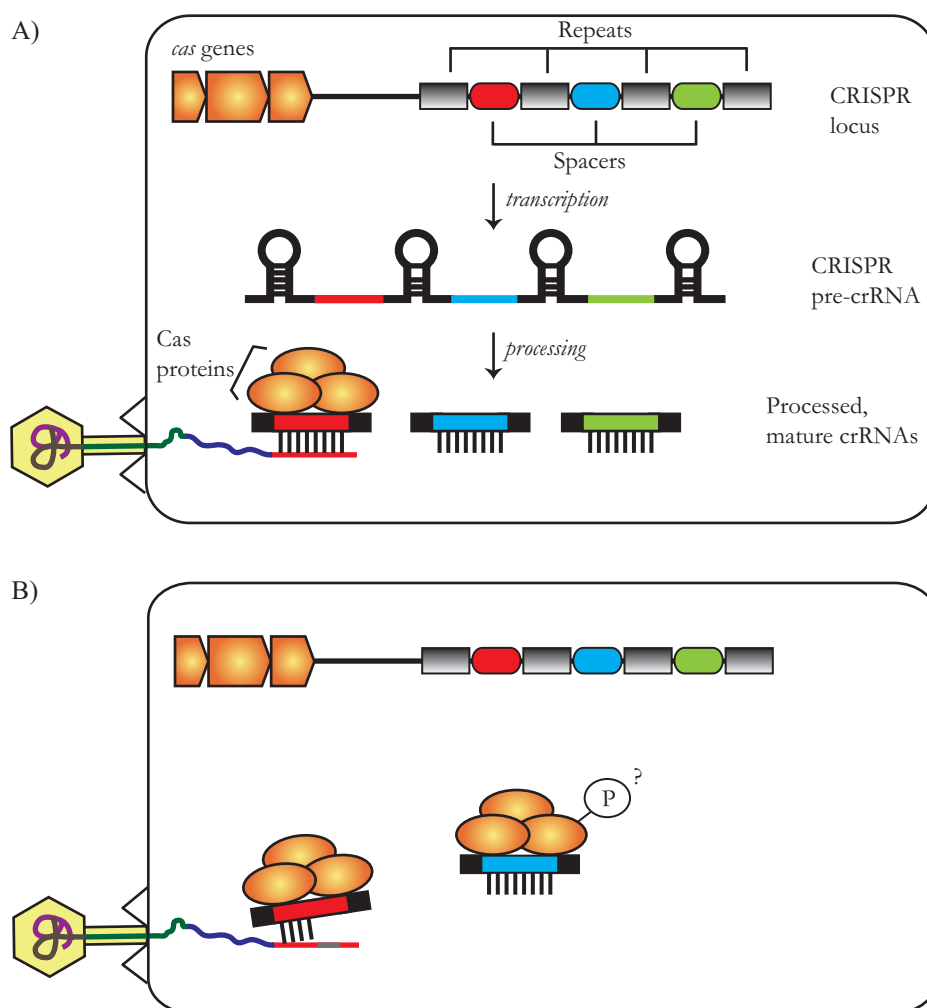


Figure 2. The CRISPR/Cas system. **A:** Mechanism of action: transcription from the repeat-spacer CRISPR locus generates a long non-coding RNA, with repeats that may sometimes assume a secondary structure. Cleavage of the repeat sequences by the Cas proteins generates crRNAs that target the phage DNA or RNA, and interfere with phage infection. **B:** Phages can evade CRISPR interference by mutation or recombination of the targeted proto-spacer sequence. Another putative evasion mechanism is phosphorylation of the Cas proteins. However, this remains to be verified.

proteins. Some Abi genes have been shown to target phage genes involved in DNA replication [45, 46], and others have been shown to target the host translation apparatus [47, 48]. For instance, in *E. coli* K-12, the Lit protease encoded by the defective ϕ 14 prophage is only activated in the presence of a short polypeptide called Gol, which is produced by the T4 phage. Once active, this protease cleaves the translation elongation factor Ef-Tu, thus leading to translational arrest and cell death (Fig. 3A). Similarly, the Prr protein (also part of a defective prophage in *E. coli* cT196) cleaves tRNA^{Lys} in response to the presence of the T4 peptide Stp [49]. Mutations in these phage peptides suppress activation of Lit or Prr and rescue the infecting phage [47, 49] (Fig. 3B).

Since Abi genes have a toxic effect on their host, they are under tight regulation [50]. In fact, many similarities can be drawn between Abi systems (and also RM systems) and toxin-antitoxin (TA) systems. TA systems are composed of a stable toxin and an unstable antitoxin [51]. Normally, the antitoxin binds and inhibits the toxin. However, a decrease in the levels of the unstable antitoxin activates the toxin and leads to growth arrest or cell death. It has been shown that this situation occurs in response to phage infection in two known TA modules, *mazEF* and *hok-sok*, which can thus cause Abi [52, 53]. Moreover, the AbiQ system was recently found to function as a protein-RNA TA pair [54], albeit exactly how this system interferes with phage replication remains unknown. Thus, it appears that some TA systems may be a subtype of Abi systems. Interestingly, RM systems may also be viewed as TA systems, since loss of the MTase (which like the antitoxin is often unstable) results in a toxic effect of the REase [55, 56].

Commonalities among phage defense systems

One of the most striking features common to all phage defense systems is their high genetic variability, which occurs as a

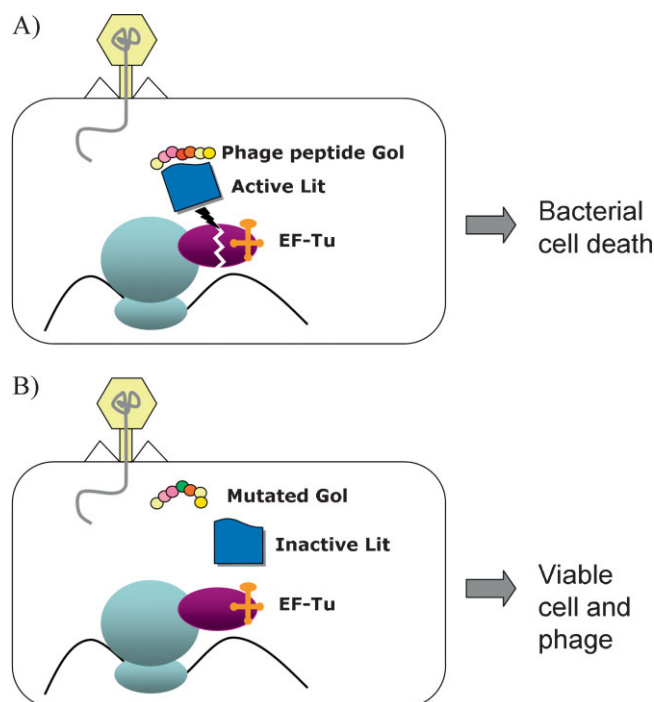


Figure 3. The Abi system. **A:** Illustration of the mechanism of the *E. coli* K-12 Abi Lit system. When the T4 phage peptide Gol is synthesized, it binds and activates the bacterial (prophage-encoded) Lit protein, which then cleaves the elongation factor EF-Tu. This leads to the arrest of protein synthesis and to bacterial cell death, with the phage trapped inside. **B:** A mutation at the Gol polypeptide reduces the activation of Lit, and rescues the phage.

consequence of the co-evolutionary arms race with phages [57]. This is manifested in the enormous numbers and types of RM systems [12, 58], of CRISPR subtypes [59, 60] and in the variety of Abi systems [44]. Moreover, the gene sequences of these systems often display a high evolutionary rate ([61–63]; A. Stern and R. Sorek, unpublished data). This variability is most easily exemplified in the diversity of RM systems: four inherently different groups of RM systems exist, classified as Type I–IV based on their subunit composition, mode of action, and cofactor requirement (reviewed in ref. 12). Each type of RM system includes an assortment of restriction enzymes that recognize different recognition sequences. For instance, nearly 4000 Type II enzymes are known of today, which are further divided into 11 overlapping subclasses [13, 64]. Interestingly, Type II REase sequences have frequently been found to be ORFan sequences, *i.e.* have no significant similarity to any other protein [65, 66]. Initially, this was seen as evidence for convergent evolution [67]. However, crystallographic studies have shown structural similarity among Type II REases, most likely representing a rapid evolutionary rate since their divergence from a common ancestor [68].

Another trait common to all defense systems presented here is their propensity to undergo LGT, sometimes between distantly related prokaryotes [69–72]. Often these systems reside on plasmids, on prophages, or on genomic islands of foreign origin, or are linked to transposase genes. This mobility allows rapid acquisition and dissemination of new systems

to counteract the invading phage, thus contributing to the extensive co-evolution with the prokaryotic parasites. However, there is an interesting alternative view, not necessarily mutually exclusive, which may explain the high variability and mobility of phage defense systems. Accordingly, these systems are selfish elements, as elaborated in the next section.

Defense systems: A burden or a benefit?

In this section, the advantages and disadvantages of microbial immune systems are reviewed. Such a discussion is incomplete without examining the pros and cons of the phage-host relationship. Evidently, for a single cell, phage predation, and killing of the host are a strong disadvantage. However, phages may also contribute directly to host fitness by supplying a pool of new and possibly beneficial genes [73]. As such, phages serve as vessels of LGT, which is one of the most important forces in microbial evolution (Box 2).

What is the associated cost of encoding phage defense mechanisms? An obvious primary cost is the energy cost linked to carrying additional genetic cargo. The fact that,

Box 2

Role of defense systems in limiting LGT

LGT is the process whereby genetic material is incorporated from a non-parental organism. There are three major mechanisms of LGT: (a) transduction by phages; (b) transformation of naked DNA; and (c) bacterial conjugation, which involves transfer of DNA *via* direct connections generated between a pair of cells. All three mechanisms involve the entry of foreign DNA into the host cell. LGT is recognized as one of the most important mechanisms of genetic innovation in both bacteria and archaea [82–84]. In fact, it is now realized that much of the genomic diversity in prokaryotes is a consequence of LGT, rather than allelic differences at the same loci [85, 86]. However, when viewed from the point of view of a single cell, LGT has a negligible chance of contributing to the fitness of the organism. Evidently, transduction by lytic phages may severely compromise the survival of a bacterial colony. Plasmids and transposons may also decrease the fitness of an organism by integrating into crucial regions of the genomes, by adding an energetic cost involved with replication of excess DNA, and by expressing harmful genes (*e.g.* refs. 87–89). Finally, even if a new beneficial gene sequence is obtained, there is only a small probability that it will integrate into the current cellular network of an organism without causing deleterious effects [90, 91]. Thus, defense mechanisms such as those described here are crucial for maintaining the genetic identity of organisms, and protecting it against the constant bombardment of potentially detrimental foreign DNA.

often, only a small portion of a given bacterial population bears a plasmid or a prophage with a given defense mechanism may be seen as evidence for this cost. A second intriguing cost of encoding a phage defense mechanism is the risk of autoimmunity. Autoimmunity, classically defined in mammalian immune systems, is the failure of the immune system to recognize what is self and what is foreign, resulting in an immune response against self. In RM systems, this is a clear danger since the restriction enzyme is often more stable than the protecting methylase [74]. Abi mechanisms are equally lethal, since errant function of the Abi gene will lead to cell death. Finally, the CRISPR system is also not “immune” to errors, and the frequent acquisition of self genetic material also leads to spacers that target the self-genome, potentially incurring autoimmunity [75], see also ref. 76).

Paradoxically, phages themselves often bear anti-phage defense systems, enabling their acquisition by host cells. To cite a few examples, the *HindIII* RM gene complex was found on a cryptic prophage in the *Haemophilus influenzae* genome [77], the *abiN* Abi gene is encoded on a prophage in *Lactococcus lactis* subsp. *cremoris* S114, and even a CRISPR array was found within a *Clostridium difficile* prophage [78]. These results initially seem counterintuitive, since what benefit is there for the phage to carry such systems? One possible explanation is that this allows superinfection exclusion, thus preventing other phages from infecting an already infected cell [11, 78, 79]. However, this phenomenon has also been viewed as evidence for the selfish nature of phage defense mechanisms.

The behavior of defense mechanisms as selfish mobile elements has been extensively discussed for the case of RM systems [55, 80], but is also applicable for many Abi systems that operate as TA systems, which also have “selfish” properties [51, 81]. The main lines of evidence in favor of this view is that (a) RM systems destroy any other invading RM system, (b) any attempt to lose the RM system will result in the death of the host, and (c) RM systems are prone to extensive mobility, and are often associated with plasmids, phages, transposons, and integrons [55]. These characteristics of the RM systems lead to an increase of their relative frequency in the bacterial population. Thus, according to this hypothesis, the defense incurred by RM systems on host cells is a mere by-product of the fact that RM systems defend themselves.

Exaptations: Alternative functions of defense systems

Intriguingly, a number of anti-phage defense systems have evolved to gain a distinct function in cellular regulation that is independent of phage restriction. Such evolutionary events were coined “exaptations” [92], a term used to describe the use of a biological structure or function for a purpose other than that for which it initially evolved. For instance, several RM systems have lost their REase activity, leaving an orphan MTase that can now take part in epigenetic modifications. The Dam methylase in *E. coli* and the CcrM methylase in *Caulobacter crescentus*, both of which originated from RM systems [93], are two such examples. Methylation by Dam has been linked to several important regulatory processes,

such as mismatch repair by the MutHLS complex, binding of the replication-initiation complex to methylated *OriC*, and regulation of bacterial pathogenicity [93]. On the other hand, the CcrM methylase has been shown to affect the cell cycle in alpha proteobacteria that encode this gene [93].

Another interesting example of exaptation of RM systems is evident in phase-variable Type III RM systems [94], whose genes can be reversibly inactivated due to tandem repeat tracts in their sequences. These repeats initiate a mechanism called slipped-strand mispairing, leading to a change in the number of repeats after DNA replication and possible frame-shift mutations [94]. Several regulatory roles have been suggested for phase-variable RM systems: (a) to allow regulated removal of the barrier against foreign DNA, thus allowing potentially beneficial uptake of DNA [95, 96], (b) autolytic self-DNA degradation or “bacterial suicide” [97–99] (discussed further below), and (c) epigenetic gene regulation via differential methylation of the genome [100]. The latter phenomenon, which allows switching different genes on and off, has been linked to pathogenicity of bacterial species by allowing colonization, immune evasion, and adaptation to novel environments [94].

TA modules, which include some Abi systems, are also known to participate in a variety of other cellular processes. For example, one of the most studied TA loci, *mazEF*, aborts translation by cleaving mRNA molecules in response to different stress signals [101, 102], one of which is phage infection [52]. An ongoing debate exists whether this action is reversible or not: reversible effects have been attributed to bacteriostatic effects, which allow reduced growth rate of each cell during nutritional stress [101, 102]. On the other hand, irreversible effects of *mazEF* have been attributed to programmed cell death that occurs in a subpopulation of cells, permitting the survival of the population of a whole [103]. Interestingly, in *Myxococcus xanthus* it was shown that the toxin MazF exists without the antitoxin MazE, and has adopted a key transcriptional regulator as an alternative antitoxin [104]. MazF mediates programmed cell death during multicellular development of this organism. To summarize, these accounts exemplify the broad evolutionary diversification of different microbial defense mechanisms, and their potential to cross boundaries from phage-encoded mechanisms, to anti-phage systems, to regulatory host mechanisms.

Conclusions

Despite our growing understanding of microbial immunity, much still remains obscure. Do defense systems work separately or in unison? What is the cost of each system? Which phages are targeted by which systems? For instance, while most characterized defense systems work against double-stranded DNA phages, RNA viruses might also be abundant [105]. Defense systems that target such viruses are yet to be discovered.

The recently discovered CRISPR system epitomizes our incomplete understanding of the complexity of bacterial defense systems. The discovery that almost half of all prokaryotes possess acquired inherited immunity came as a surprise to the scientific community, given our initial tendency to view prokaryotes as less complex organisms. However, since it is

now realized that prokaryotes are faced with a constant threat of predation, it is becoming clearer that the microbial immune system must be highly complex. Nevertheless, the CRISPR system and the Abi systems have a highly sporadic distribution across the prokaryotic phylogeny (Table 1). Furthermore, while RM systems are present in ~90% of prokaryotic species [13], this system is far from infallible, since it has been shown that phages have a non-negligible probability of escaping the REase and being methylated by the MTase. [106]. Thus, all three mechanisms reviewed here are either somewhat leaky, or span only part of the phylogenetic breadth of bacteria and archaea. Given the intensive arms race between bacteria and phages, it seems probable that there are other yet unknown defense mechanisms waiting to be revealed.

So how does one set about the search for novel prokaryotic defense systems, and how does one learn more about existing ones? It is likely that many solutions to these questions will be made possible, thanks to advances in sequencing technologies such as whole-transcriptome studies and metagenomics. For instance, a seminal work by Andersson and Banfield [42] reconstructed viral and host population genomes from community genomic data, and showed how CRISPR spacers correlate with the location of coexisting viruses and hosts. In addition, controlled experimental evolution studies, such as those performed in *Pseudomonas fluorescens* and its phage, enable a direct survey of the changes that occur in the phage and host population with and without infection [107–109]. These studies are also likely to shed light on current defense systems and possibly discover novel ones.

More specifically, the search for novel defense mechanisms may be based on the well-established common properties of both eukaryotic and prokaryotic immune systems: their high rate of evolution, and their tendency to undergo LGT. Thus, it may be that a reservoir of novel defense mechanisms lies in the most variable regions of bacterial and archaeal genomes, known as genomic islands [110]. A strategy that focuses on such islands to search for novel phage resistance mechanisms might lead, in the future, to surprising discoveries.

Acknowledgments

R.S. is an EMBO Young Investigator. He was supported, in part, by the ISF-FIRST program (grant 1615/09), NIH R01AI082376-01, ERC-StG, the Wolfson Family Trust miRNA research program, the Minerva foundation, and the Yeda-Sela Center for basic research. A.S. was supported by the Clore post-doctoral fellowship.

References

- Torrella F, Morita RY. 1979. Evidence by electron micrographs for a high incidence of bacteriophage particles in the waters of Yaquina Bay, Oregon: ecological and taxonomical implications. *Appl Environ Microbiol* **37**: 774–8.
- Riesenfeld CS, Schloss PD, Handelsman J. 2004. Metagenomics: genomic analysis of microbial communities. *Annu Rev Genet* **38**: 525–52.
- Bergh O, Borsheim KY, Bratbak G, Haldal M. 1989. High abundance of viruses found in aquatic environments. *Nature* **340**: 467–8.
- Chibani-Chennoufi S, Bruttin A, Dillmann ML, Brussow H. 2004. Phage-host interaction: an ecological perspective. *J Bacteriol* **186**: 3677.
- Wommack KE, Colwell RR. 2000. Virioplankton: viruses in aquatic ecosystems. *Microbiol Mol Biol Rev* **64**: 69.
- Fuhrman JA. 1999. Marine viruses and their biogeochemical and ecological effects. *Nature* **399**: 541–8.
- Suttle CA. 2007. Marine viruses – major players in the global ecosystem. *Nat Rev Microbiol* **5**: 801–12.
- Comeau AM, Krisch HM. 2005. War is peace – dispatches from the bacterial and phage killing fields. *Curr Opin Microbiol* **8**: 488–94.
- Brussow H, Canchaya C, Hardt WD. 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev* **68**: 560–602.
- Van Valen L. 1973. A new evolutionary law. *Evol Theor* **1**: 1–30.
- Labrie SJ, Samson JE, Moineau S. 2010. Bacteriophage resistance mechanisms. *Nat Rev Microbiol* **8**: 317–27.
- Tock MR, Dryden DT. 2005. The biology of restriction and anti-restriction. *Curr Opin Microbiol* **8**: 466–72.
- Roberts RJ, Vincze T, Posfai J, Macelis D. 2010. REBASE – a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res* **38**: D234–6.
- Roberts RJ, Vincze T, Posfai J, Macelis D. 2005. REBASE – restriction enzymes and DNA methyltransferases. *Nucleic Acids Res* **33**: D230.
- Kruger DH, Bickle TA. 1983. Bacteriophage survival: multiple mechanisms for avoiding the deoxyribonucleic acid restriction systems of their hosts. *Microbiol Rev* **47**: 345–60.
- Bandyopadhyay PK, Studier FW, Hamilton DL, Yuan R. 1985. Inhibition of the type I restriction-modification enzymes EcoB and EcoK by the gene 0.3 protein of bacteriophage T7. *J Mol Biol* **182**: 567–78.
- Wang Z, Mosbaugh DW. 1989. Uracil-DNA glycosylase inhibitor gene of bacteriophage PBS2 encodes a binding protein specific for uracil-DNA glycosylase. *J Biol Chem* **264**: 1163–71.
- Wang Z, Mosbaugh DW. 1988. Uracil-DNA glycosylase inhibitor of bacteriophage PBS2: cloning and effects of expression of the inhibitor gene in *Escherichia coli*. *J Bacteriol* **170**: 1082–91.
- Putnam CD, Tainer JA. 2005. Protein mimicry of DNA and pathway regulation. *DNA Repair (Amst)* **4**: 1410–20.
- Rocha EP, Danchin A, Viari A. 2001. Evolutionary role of restriction/modification systems as revealed by comparative genome analysis. *Genome Res* **11**: 946–58.
- Raleigh EA, Wilson G. 1986. *Escherichia coli* K-12 restricts DNA containing 5-methylcytosine. *Proc Natl Acad Sci USA* **83**: 9070–4.
- Bair CL, Rifat D, Black LW. 2007. Exclusion of glucosyl-hydroxymethylcytosine DNA containing bacteriophages is overcome by the injected protein inhibitor IPI*. *J Mol Biol* **366**: 779–89.
- Bair CL, Black LW. 2007. A type IV modification dependent restriction nuclease that targets glucosylated hydroxymethyl cytosine modified DNAs. *J Mol Biol* **366**: 768–78.
- Rifat D, Wright NT, Varney KM, Weber DJ, et al. 2008. Restriction endonuclease inhibitor IPI* of bacteriophage T4: a novel structure for a dedicated target. *J Mol Biol* **375**: 720–34.
- Ishino Y, Shinagawa H, Makino K, Amemura M, et al. 1987. Nucleotide sequence of the lap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol* **169**: 5429.
- Jansen R, van Embden JDA, Gaastra W, Schouls LM. 2002. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* **43**: 1565–75.
- Sorek R, Kunin V, Hugenholtz P. 2008. CRISPR – a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat Rev Microbiol* **6**: 181–6.
- Grissa I, Vergnaud G, Pourcel C. 2007. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* **8**: 172.
- Bolotin A, Quinquis B, Sorokin A, Ehrlich SD. 2005. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**: 2551–61.
- Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E. 2005. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* **60**: 174–82.
- Pourcel C, Salvignol G, Vergnaud G. 2005. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**: 653–63.

32. Barrangou R, Fremaux C, Deveau H, Richards M, et al. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**: 1709–12.
33. Horvath P, Barrangou R. 2010. CRISPR/Cas, the immune system of bacteria and archaea. *Science* **327**: 167–70.
34. Marraffini LA, Sontheimer EJ. 2010. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet* **11**: 181–90.
35. van der Oost J, Jore MM, Westra ER, Lundgren M, et al. 2009. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem Sci* **34**: 401–7.
36. Marraffini LA, Sontheimer EJ. 2008. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322**: 1843.
37. Brouns SJ, Jore MM, Lundgren M, Westra ER, et al. 2008. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**: 960–4.
38. Makarova KS, Grishin NV, Shabalina SA, Wolf YI, et al. 2006. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol Direct* **1**: 7.
39. Hale CR, Zhao P, Olson S, Duff MO, et al. 2009. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**: 945–56.
40. Deveau H, Barrangou R, Garneau JE, Labonte J, et al. 2008. Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J Bacteriol* **190**: 1390–400.
41. Heidelberg JF, Nelson WC, Schoenfeld T, Bhaya D. 2009. Germ warfare in a microbial mat community: CRISPRs provide insights into the co-evolution of host and viral genomes. *PLoS One* **4**: e4169.
42. Andersson AF, Banfield JF. 2008. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* **320**: 1047–50.
43. Qimron U, Tabor S, Richardson CC. 2010. New details about bacteriophage T7-host interactions. *Microbe* **5**: 117–20.
44. Chopin MC, Chopin A, Bidnenko E. 2005. Phage abortive infection in lactococci: variations on a theme. *Curr Opin Microbiol* **8**: 473–9.
45. Durmaz E, Klaenhammer TR. 2007. Abortive phage resistance mechanism AbiZ speeds the lysis clock to cause premature lysis of phage-infected *Lactococcus lactis*. *J Bacteriol* **189**: 1417.
46. Bidnenko E, Ehrlich D, Chopin MC. 1995. Phage operon involved in sensitivity to the *Lactococcus lactis* abortive infection mechanism AbiD1. *J Bacteriol* **177**: 3824.
47. Georgiou T, Yu YTN, Ekinwe S, Buttner MJ, et al. Specific peptide-activated proteolytic cleavage of *Escherichia coli* elongation factor Tu. 1998. *Proc Natl Acad Sci USA* **95**: 2891.
48. Morad I, Chapman-Shimshoni D, Amitur M, Kaufmann G. 1993. Functional expression and properties of the tRNA (Lys)-specific core anticodon nuclease encoded by *Escherichia coli* prC. *J Biol Chem* **268**: 26842.
49. Snyder L. 1995. Phage-exclusion enzymes: a bonanza of biochemical and cell biology reagents? *Mol Microbiol* **15**: 415–20.
50. Blower TR, Fineran PC, Johnson MJ, Toth IK, et al. 2009. Mutagenesis and functional characterization of the RNA and protein components of the toxIN abortive infection and toxin-antitoxin locus of *Erwinia*. *J Bacteriol* **191**: 6029–39.
51. Gerdes K, Christensen SK, Lobner-Olesen A. 2005. Prokaryotic toxin-antitoxin stress response loci. *Nat Rev Microbiol* **3**: 371–82.
52. Hazan R, Engelberg-Kulka H. 2004. *Escherichia coli* mazEF-mediated cell death as a defense mechanism that inhibits the spread of phage P1. *Mol Genet Genomics* **272**: 227–34.
53. Pecota DC, Wood TK. 1996. Exclusion of T4 phage by the hok/sok killer locus from plasmid R1. *J Bacteriol* **178**: 2044.
54. Fineran PC, Blower TR, Foulds IJ, Humphreys DP, et al. 2009. The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. *Proc Natl Acad Sci USA* **106**: 894–9.
55. Kobayashi I. 2001. Behavior of restriction-modification systems as selfish mobile elements and their impact on genome evolution. *Nucleic Acids Res* **29**: 3742–56.
56. Hayes F. 2003. Toxins-antitoxins: plasmid maintenance, programmed cell death, and cell cycle arrest. *Science* **301**: 1496–9.
57. Hoskisson PA, Smith MC. 2007. Hypervariation and phase variation in the bacteriophage 'resistome'. *Curr Opin Microbiol* **10**: 396–400.
58. Orlowski J, Bujnicki JM. 2008. Structural and evolutionary classification of Type II restriction enzymes based on theoretical and experimental analyses. *Nucleic Acids Res* **36**: 3552.
59. Haft DH, Selengut J, Mongodin EF, Nelson KE. 2005. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput Biol* **1**: e60.
60. Kunin V, Sorek R, Hugenoltz P. 2007. Evolutionary conservation of sequence and secondary structures in CRISPR repeats. *Genome Biol* **8**: R61.
61. Sharp PM, Kelleher JE, Daniel AS, Cowan GM, et al. 1992. Roles of selection and recombination in the evolution of type I restriction-modification systems in enterobacteria. *Proc Natl Acad Sci USA* **89**: 9836–40.
62. Murray NE, Daniel AS, Cowan GM, Sharp PM. 1993. Conservation of motifs within the unusually variable polypeptide sequences of type I restriction and modification enzymes. *Mol Microbiol* **9**: 133–43.
63. Zheng Y, Roberts RJ, Kasif S. 2004. Identification of genes with fast-evolving regions in microbial genomes. *Nucleic Acids Res* **32**: 6347–57.
64. Roberts RJ, Belfort M, Bestor T, Bhagwat AS, et al. 2003. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res* **31**: 1805–12.
65. Kroger M, Hobom G, Schutte H, Mayer H. 1984. Eight new restriction endonucleases from *Herpetosiphon giganteus*-divergent evolution in a family of enzymes. *Nucleic Acids Res* **12**: 3127.
66. Mullings R, Bennett SP, Brown NL. 1988. Investigation of sequence homology in a group of type-II restriction/modification isoschizomers. *Gene* **74**: 245–51.
67. Wilson G, Murray NE. 1991. Restriction and modification systems. *Annu Rev Genet* **25**: 585–627.
68. Venclovas C, Timinskas A, Siksnys V. 1994. Five-stranded beta-sheet sandwiched with two alpha-helices: a structural link between restriction endonucleases *EcoRI* and *EcoRV*. *Proteins* **20**: 279–82.
69. Nelson KE, Clayton RA, Gill SR, Gwinn ML, et al. 1999. Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**: 323–9.
70. Jeltsch A, Pingoud A. 1996. Horizontal gene transfer contributes to the wide distribution and evolution of type II restriction-modification systems. *J Mol Evol* **42**: 91–6.
71. Haaber J, Moineau S, Hammer K. 2009. Activation and transfer of the chromosomal phage resistance mechanism AbiV in *Lactococcus lactis*. *Appl Environ Microbiol* **75**: 3358–61.
72. Godde JS, Bickerton A. 2006. The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *J Mol Evol* **62**: 718–29.
73. Gogarten JP, Townsend JP. 2005. Horizontal gene transfer, genome innovation and evolution. *Nat Rev Microbiol* **3**: 679–87.
74. Naito T, Kusano K, Kobayashi I. 1995. Selfish behavior of restriction-modification systems. *Science* **267**: 897–9.
75. Stern A, Keren L, Wurtzel O, Amitai G, et al. 2010. Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet* **26**: 335–40.
76. Marraffini LA, Sontheimer EJ. 2010. Self versus non-self discrimination during CRISPR RNA-directed immunity. *Nature* **463**: 568–71.
77. Hendrix RW, Smith MC, Burns RN, Ford ME, et al. 1999. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc Natl Acad Sci USA* **96**: 2192–7.
78. Sebahia M, Wren BW, Mullany P, Fairweather NF, et al. 2006. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nat Genet* **38**: 779–86.
79. Lu MJ, Henning U. 1994. Superinfection exclusion by T-even-type coliphages. *Trends Microbiol* **2**: 137–9.
80. Nakayama Y, Kobayashi I. 1998. Restriction-modification gene complexes as selfish gene entities: roles of a regulatory system in their establishment, maintenance, and apoptotic mutual exclusion. *Proc Natl Acad Sci USA* **95**: 6442–7.
81. Makarova KS, Wolf YI, Koonin EV. 2009. Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol Direct* **4**: 19.
82. Gogarten JP, Doolittle WF, Lawrence JG. 2002. Prokaryotic evolution in light of gene transfer. *Mol Biol Evol* **19**: 2226–38.
83. Koonin EV, Makarova KS, Aravind L. 2001. Horizontal gene transfer in prokaryotes: quantification and classification. *Annu Rev Microbiol* **55**: 709–42.
84. Dagan T, Artzy-Randrup Y, Martin W. 2008. Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. *Proc Natl Acad Sci USA* **105**: 10039–44.
85. Whitaker RJ, Banfield JF. 2006. Population genomics in natural microbial communities. *Trends Ecol Evol* **21**: 508–16.
86. Medini D, Serruto D, Parkhill J, Relman DA, et al. 2008. Microbiology in the post-genomic era. *Nat Rev Microbiol* **6**: 419–30.
87. Modi RI, Adams J. 1991. Coevolution in bacterial-plasmid populations. *Evolution* **45**: 656–67.

88. Zund P, Lebek G. 1980. Generation time-prolonging R plasmids: correlation between increases in the generation time of *Escherichia coli* caused by R plasmids and their molecular size. *Plasmid* **3**: 65–9.
89. Bouma JE, Lenski RE. 1988. Evolution of a bacteria/plasmid association. *Nature* **335**: 351–2.
90. Kedzierska B, Glinkowska M, Iwanicki A, Obuchowski M, et al. 2003. Toxicity of the bacteriophage lambda cII gene product to *Escherichia coli* arises from inhibition of host cell DNA replication. *Virology* **313**: 622–8.
91. Lercher MJ, Pal C. 2008. Integration of horizontally transferred genes into regulatory interaction networks takes many million years. *Mol Biol Evol* **25**: 559–67.
92. Brosius J, Gould SJ. 1992. On “genomenclature”: a comprehensive (and respectful) taxonomy for pseudogenes and other “junk DNA”. *Proc Natl Acad Sci USA* **89**: 10706–10.
93. Marinus MG, Casadesus J. 2009. Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more. *FEMS Microbiol Rev* **33**: 488–503.
94. Srikhanta YN, Maguire TL, Stacey KJ, Grimmond SM, et al. 2005. The phasevarion: a genetic system controlling coordinated, random switching of expression of multiple genes. *Proc Natl Acad Sci USA* **102**: 5547–51.
95. Ando T, Xu Q, Torres M, Kusugami K, et al. 2000. Restriction-modification system differences in *Helicobacter pylori* are a barrier to inter-strain plasmid transfer. *Mol Microbiol* **37**: 1052–65.
96. Donahue JP, Israel DA, Peek RM, Blaser MJ, et al. 2000. Overcoming the restriction barrier to plasmid transformation of *Helicobacter pylori*. *Mol Microbiol* **37**: 1066–74.
97. Dybvig K, Sitaraman R, French CT. 1998. A family of phase-variable restriction enzymes with differing specificities generated by high-frequency gene rearrangements. *Proc Natl Acad Sci USA* **95**: 13923–8.
98. Saunders NJ, Peden JF, Hood DW, Moxon ER. 1998. Simple sequence repeats in the *Helicobacter pylori* genome. *Mol Microbiol* **27**: 1091–8.
99. Hamilton HL, Dillard JP. 2006. Natural transformation of *Neisseria gonorrhoeae*: from DNA donation to homologous recombination. *Mol Microbiol* **59**: 376–85.
100. Seib KL, Peak IR, Jennings MP. 2002. Phase variable restriction-modification systems in *Moraxella catarrhalis*. *FEMS Immunol Med Microbiol* **32**: 159–65.
101. Pedersen K, Christensen SK, Gerdes K. 2002. Rapid induction and reversal of a bacteriostatic condition by controlled expression of toxins and antitoxins. *Mol Microbiol* **45**: 501–10.
102. Christensen SK, Pedersen K, Hansen FG, Gerdes K. 2003. Toxin-antitoxin loci as stress-response-elements: ChpAK/MazF and ChpBK cleave translated RNAs and are counteracted by tmRNA. *J Mol Biol* **332**: 809–19.
103. Engelberg-Kulka H, Amitai S, Kolodkin-Gal I, Hazan R. 2006. Bacterial programmed cell death and multicellular behavior in bacteria. *PLoS Genet* **2**: e135.
104. Nariya H, Inouye M. 2008. MazF, an mRNA interferase, mediates programmed cell death during multicellular *Myxococcus* development. *Cell* **132**: 55–66.
105. Lang AS, Rise ML, Culley AI, Steward GF. 2009. RNA viruses in the sea. *FEMS Microbiol Rev* **33**: 295–323.
106. Korona R, Levin BR. 1993. Phage-mediated selection and the evolution and maintenance of restriction-modification. *Evolution* **47**: 556–75.
107. Paterson S, Vogwill T, Buckling A, Benmayor R, et al. 2010. Antagonistic coevolution accelerates molecular evolution. *Nature* **464**: 275–8.
108. Buckling A, Wei Y, Massey RC, Brockhurst MA, et al. Antagonistic coevolution with parasites increases the cost of host deleterious mutations. *Proc Biol Sci* 2006. **273**: 45–9.
109. Wichman HA, Badgett MR, Scott LA, Boulianne CM, et al. 1999. Different trajectories of parallel evolution during viral adaptation. *Science* **285**: 422–4.
110. Juhas M, van der Meer JR, Gaillard M, Harding RM, et al. 2009. Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol Rev* **33**: 376–93.
111. Pandey DP, Gerdes K. 2005. Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. *Nucleic Acids Res* **33**: 966–76.
112. Iida S, Streiff MB, Bickle TA, Arber W. 1987. Two DNA antirestriction systems of bacteriophage P1, darA, and darB: characterization of darA-phages. *Virology* **157**: 156–66.
113. Dryden DT, Tock MR. 2006. DNA mimicry by proteins. *Biochem Soc Trans* **34**: 317–9.